EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Chemical Similarity and Threshold of Toxicological Concern (TTC) Approaches
Report of an ECB Workshop held in Ispra, November 2005

Grace Patlewicz, Ana Gallegos Saliner, Manuela Pavan, Andrew Worth, Romualdo Benigni, Aynur Aptula, Arianna Bassan, Cecilia Bossa, Agneta Falk-Filipsson, Val Gillet, Nina Jeliazkova, Andrew McDougal, Jordi Mestres, Ian Munro, Tatiana Netzeva, Bob Safford, Brigitte Simon-Hettich, Ivanka Tsakovska, Maria Wallén, Chihae Yang

IHCP

2007

EUR 22657 EN

The mission of IHCP is to provide scientific support to the development and implementation of EU policies related to health and consumer protection.
The IHCP carries out research to improve the understanding of potential health risks posed by chemicals, physical and biological agents from various sources to which consumers are exposed.

- As a Research Institute, the IHCP contributes to the improvement of scientific knowledge on health care methods and consumer issues.
- As a European Institution, the IHCP participates, at an international level and in collaboration with Member States Authorities, in R&D and regulatory actions intended to improve consumer tutelage.
- As a European Commission Service, the IHCP acts as an independent advisor for the implementation of risk assessment, monitoring and validation of procedures aimed to insure EU citizens of the use of health and consumer services or products.

European Commission
Directorate-General Joint Research Centre
Institute IHCP

Contact information Andrew Worth
Address: European Chemicals Bureau, TP81
E-mail: Andrew.Worth@jrc.it
Tel.: +39 0332 789566
Fax: +39 0332 786717

http://ihcp.jrc.cec.eu.int/; http://ecb.jrc.it/
http://www.jrc.cec.eu.int

EUR 22657 EN

ISSN 1018-5593

Luxembourg: Office for Official Publications of the European Communities

Printed in Italy

Chemical Similarity and Threshold of Toxicological Concern (TTC) Approaches

Report of an ECB Workshop held in Ispra, November 2005

Grace Patlewicz[1], Ana Gallegos Saliner[1], Manuela Pavan[1], Romualdo Benigni[2],Andrew Worth[1], Aynur Aptula[3], Arianna Bassan[1], Cecilia Bossa[2], Agneta Falk-Filipsson[4], Val Gillet[5], Nina Jeliazkova[6], Andrew McDougal[7], Jordi Mestres[8], Ian Munro[9], Tatiana Netzeva[1], Bob Safford[3], Brigitte Simon-Hettich[10], Ivanka Tsakovska[1], Maria Wallén[4], Chihae Yang[11]

[1]European Chemicals Bureau TP 582, Institute of Health and Consumer Protection, Joint Research Centre, European Commission, Ispra 21020, ITALY

[2]Experimental and Computational Carcinogenesis Unit, Environment and Health Department, Dipartimento di Ambiente e connessa Prevenzione, IstitutoSuperiore di Sanita, Rome, ITALY

[3]Safety & Environmental Assurance Centre, Unilever R&D, Colworth House, Sharnbrook, Bedford MK44 1LQ, UK

[4]KEMI, Swedish Chemical Inspectorate, SE - 172 13 Sundbyberg, Stockholm, SWEDEN

[5]Computational Informatics Research Group, Department of Information Studies, University of Sheffield, UK

[6]IDEA Consult Ltd, Sofia, BULGARIA

[7]Division of Food Contact Notifications, Office of Food Additive Safety, FDA, Washington, USA

[8]Research Unit on Biomedical Informatics, Municipal Institute of Medical Research, Barcelona, SPAIN

[9]CANTOX Health Sciences International, Ontario, CANADA

[10]Institute of Toxicology, Merck, Darmstadt, GERMANY

[11]Leadscope® Inc, Columbus, Ohio, USA

Chemical Similarity and Threshold of Toxicological Concern (TTC) Approaches

Report of an ECB Workshop held in Ispra, November 2005

LIST OF ABBREVIATIONS

ECB	European Chemicals Bureau
GPMT	Guinea Pig Maximisation Test
LLNA	Local Lymph Node Assay
PCA	Principal Components Analysis
(Q)SAR	(Quantitative) Structure-Activity Relationship
RAI	Relative Alkylation Index
REACH	Registration, Evaluation and Authorisation of CHemicals
RP	Recursive partitioning
SOM	Self Organising Map
TTC	Thresholds of Toxicological Concern

1. Background

Chemical similarity is a widely used concept in toxicology, and is based on the hypothesis that similar compounds should have similar biological activities. This forms the underlying basis for developing structurally based Threshold of Toxicological Concern (TTC), performing read across, forming chemical categories or developing (Quantitative) Structure-Activity Relationships ((Q)SARs).

Chemical similarity is often perceived as structural similarity but in fact there are a number of other approaches that can be used to assess similarity. A similarity analysis usually comprises two main steps. Firstly the chemical structures to be compared are characterised in terms of descriptors which encode their physicochemical, topological, geometrical and/or surface properties. A second step involves a quantitative comparison of those descriptors using similarity (or dissimilarity) indices.

Note: Chemical similarity is a relative concept, two chemicals are only similar with respect to a defined parameter. Hence as an example two chemicals may behave similarly with respect to their mutagenicity profile.

Under the REACH (Registration, Evaluation, and Authorisation of CHemicals) [1] legislation, non-testing approaches such as QSARs, SARs, read-across and chemical categories, are foreseen to be more widely used. Guidance on the use of (Q)SARs is provided in Annex XI of the proposed REACH legislation [2]. It states that (Q)SARs may be used to indicate the presence or absence of a certain dangerous property if the following conditions are met:

- results are derived from a (Q)SAR model whose scientific validity has been established
- the substance falls within the applicability domain of the (Q)SAR model
- results are adequate for the purpose of classification and labelling and/or risk assessment, and
- adequate and reliable documentation of the applied method is provided

Annex XI also states that chemicals may be classified on the basis of their (eco)toxicological hazard by applying chemical grouping approaches (e.g. read-across, chemical categories).

To date, the acceptance of (Q)SARs has been limited due partly to a lack of common understanding in how to evaluate the scientific validity of the models. In November 2004 the OECD Member Countries and the Commission agreed five 'OECD principles for (Q)SAR validation' [3]. The principles provide a useful framework and practical guidance on how to demonstrate concordance for a (Q)SAR has been developed [4]. Limited guidance for developing chemical categories does exist [5] but the tools for their practical implementation are still lacking [6]. Work is underway to develop practical non-prescriptive guidance on how non-testing approaches can be applied in practice (as part of one of the REACH Implementation

Projects, specifically RIP3.3-2), and to develop software tools to make it easier to establish, justify, and document read-across, chemical category and (Q)SAR approaches. This guidance should explain and illustrate:

- The commonalities and differences between SARs, read-across and categories
- How to justify and report qualitative and quantitative read-across (in terms of supporting information)
- How to build a category (practical details), including examples of qualitative and quantitative read-across
- How to evaluate the robustness and applicability domain of a category or QSAR model
- How to justify and report a category proposal or a QSAR model prediction, in terms of its underlying rationale, the scientific basis and validity, including a thorough analysis whether the predictions are within the applicability domain for the category/model or not.

Thus, chemical category development, structural TTC, and read across are all dependent on grouping chemicals on the basis of their chemical similarity. The challenge therefore is how best to encode "similarity in activity" in a meaningful way that facilitates the use of these approaches.

To partially address this need, an ECB workshop entitled "Chemical Similarity and TTC approaches" was organised with thirteen external experts in the area of TTC and chemical similarity to discuss the underlying terminology and concepts, and review existing approaches for the grouping of chemicals both for the development of thresholds (TTC) and chemical categories. This report summarises the outcomes and recommendations of this workshop.

The aims of this workshop were twofold:

- to discuss the underlying terminology and concepts, and review existing approaches for the grouping of chemicals both for the development of thresholds (TTC) and chemical categories.
- to communicate the output of recent work that had been undertaken on TTC and chemical similarity within ECB.

The outcome of the workshop was intended to guide future research efforts in this area that could potentially be used to support the development of specific EU and/or OECD guidance documentation on grouping methods.

REACH Context

The development of categories is envisaged to bring together companies to propose initial groupings of substances. This would involve a collation of available test data for those substances to build up a data

matrix to enable trends and commonalities to be identified. The next step in this process would involve a more detailed assessment which might well result in the formation of sub-categories where trends and relationships are found to only hold for a subset of substances within that group or for only one or more endpoints.

At this point, the use of alternative approaches and tools (such as the QSAR techniques, chemical similarity indices or analogues) can help to refine the grouping proposed. Indeed the use of some of the techniques might equally be applicable within the context of an individual company's inventory in order to establish initial groupings (a category hypotheses) of similar substances.

The objective of the workshop was not to consider the administrative procedure within REACH and the implications thereof but simply to explore the technical and scientific solutions that might assist in developing groupings. Therefore this report does not address the potential mapping of scientific considerations on to the administrative procedures of REACH.

2.Discussion outcomes

A number of key questions were brainstormed during the course of the workshop including:

a) Can chemical similarity play a role? What indices and approaches are available and what are their strengths and limitations?
b) Can chemical similarity indices be used as a means of comparing 2 or more chemicals in terms of their biological activity?
c) Can a flowchart or a process map be formulated to provide a systematic approach for the development of chemical categories and read across?

Overall, there was consensus that similarity indices have the potential to provide useful information to relate the biological similarity of two or more chemicals but that this was just one aspect in a larger process.

The group also agreed that similarity was context dependent and that the concept of an endpoint specific category was more appropriate when investigating potential trends and correlations between members of a group.

Much of the discussion during the course of the meeting focussed on identifying the key components that would form the basis of a flowchart (or integrated process map) for describing the practical steps in developing chemical categories/groupings. A flowchart was constructed to try to capture the work flow involved in formulating groupings (see the Flowchart as outlined in Scheme 2).

3.Grouping Methods

Grouping methods may be nominally categorised into one of four classes – knowledge-based, analogue-based, unsupervised, and supervised. The purpose or objective of the grouping will also play a role as to which approach(es) might be most appropriate to use in a given context and depending on the information available. Table 1 provides a summary of the plausible situations and their respective data requirements.

3.1 Knowledge-based approaches

Knowledge-based approaches encompass “human expert rules”. Human experience for specific endpoints is often and conveniently encoded in the form of structural alerts. This type of approach may be used in either a top-down up mode where the structural rules are used as seeds to form smaller groupings from a starting inventory/dataset of chemicals or alternatively in a bottom up approach where the alerts are supplemented with other structural analogues (see later for a more detailed description of analogue based approaches) to build up groupings. To illustrate the sorts of human expert rules available, two specific endpoints, skin sensitisation and carcinogenicity are discussed as examples.

3.1.1 Skin sensitisation

Skin sensitisation results from a T-lymphocyte mediated immune response to a chemical allergen that comes into contact with the skin. A chemical penetrates the skin and binds to a carrier protein typically by a covalent bond to form an antigenic hapten-protein complex. This complex is processed by antigen presenting cells principally dendritic Langerhans cells (LC) of the epidermis. These cells then migrate to the draining lymph node where they present the chemical to T-lymphocyte to provide the stimulus for antigen-specific commitment and the production of memory and effector T lymphocytes. Subsequent contact with a sufficient dose of the chemical will then result in the expression of the clinical signs of allergic contact dermatitis (ACD). Thus chemicals need to overcome a number of hurdles in order to induce skin sensitisation. These comprise:

- Penetration into the viable epidermis across the *stratum corneum*.
- Formation of a stable association with protein to create an immunogenic complex. This requires that a chemical is inherently protein-reactive, or can be transformed chemically or metabolically to a protein-reactive species. Typically the stable association is thought to be a covalent one.
- Delivery of a dermal trauma sufficient to induce and up-regulate those epidermal cytokines that are necessary for the mobilisation, migration and maturation of LC.
- Being inherently immunogenic such that a T lymphocyte response of sufficient magnitude is stimulated.

If these hurdles are not successfully overcome then skin sensitisation will either not occur or will be sub-optimal [7-10].

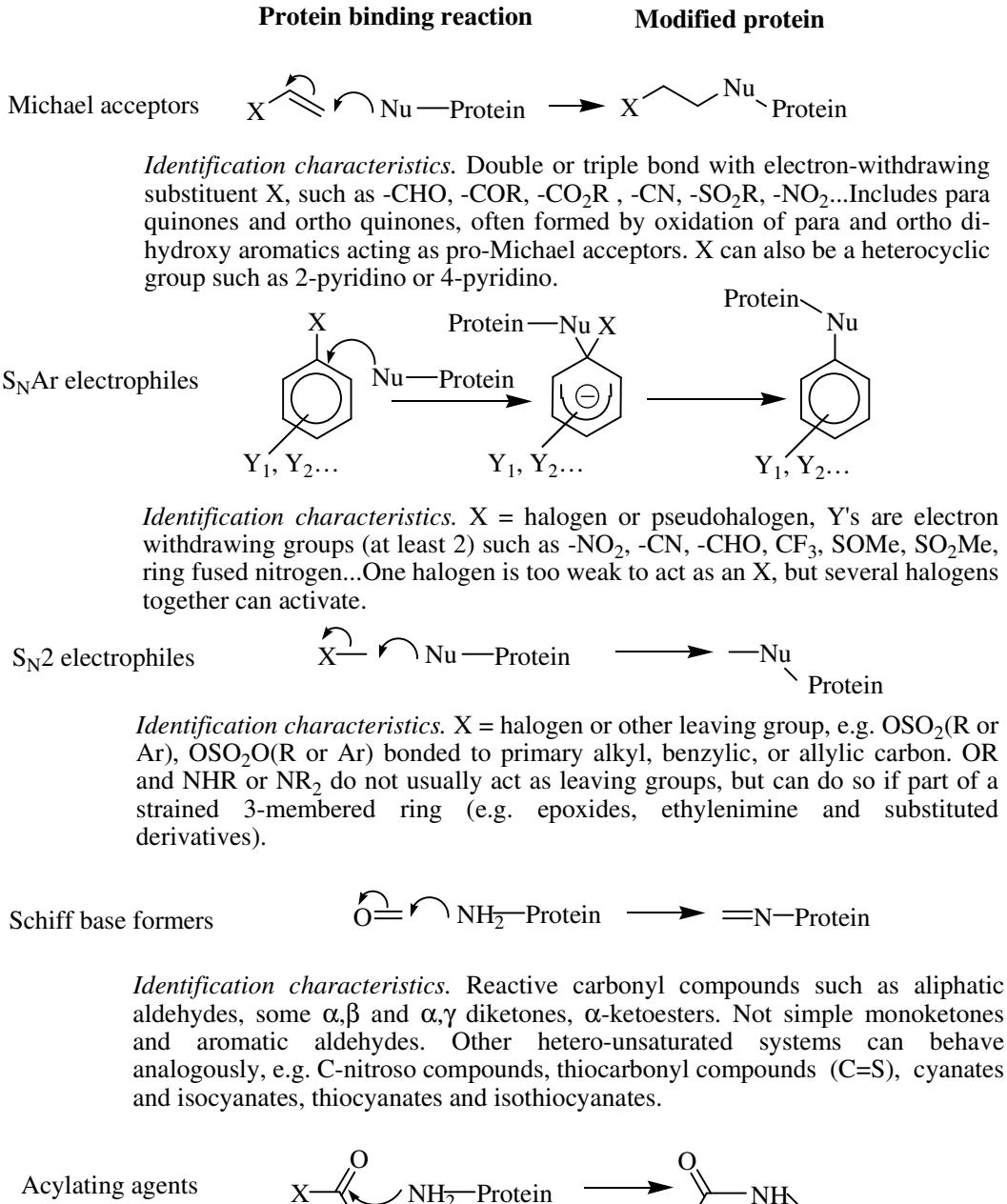
A key step is the formation of a stable association (usually covalent) with a protein. Landsteiner and Jacobs [11] first compared the results of guinea pig sensitisation on a series of aromatic halides and pseudohalides with the results of experiments in which reactions with aniline (as the model protein) were attempted to establish the connection between the ability of chemicals to react with proteins to form covalently linked conjugates and their sensitisation potential. This spurned the belief that is still withstanding today that skin sensitisation is underpinned by mechanisms based on chemical reactivity. This also explains why so many of the QSARs developed to date have exploited the relationship of covalent reactions between a model nucleophile and the chemical itself. The chemical behaves as an electrophile whereas the protein behaves as a nucleophile. There are various types of electrophile-nucleophile reactions encountered in skin sensitisation, perhaps the most common are: Michael-type reactions; S_N2 reactions; S_NAr reactions; acylation reactions and Schiff-base formation.

For direct acting electrophiles, variation of skin sensitisation potential and dose-response relationships within a chemical class has been found to be effectively modelled using the Relative Alkylation Index (RAI)

approach, a mechanism-based model developed by Roberts and Williams [12]. This model has been used to study a number of structural classes of skin sensitising chemicals. Originally the model was applied to guinea pig test data (guinea pig maximisation tests; GPMTs). The introduction of the LLNA (Local Lymph Node Assay) with its quantitative endpoint for skin sensitising potency provided a step change in capability to skin sensitisation (Q)SARs to relate chemicals of not just the same chemical class but the same mechanism of action. In this way, chemicals such as 1,2-diketones and certain aldehydes have been related together as Schiff base formers [13]. There have been other instances of chemicals that fall within specific applicability domains that are related to chemical mechanisms which is a more meaningful representation of chemical similarity than structural similarities. Examples of on-going work include [14-17]. In [14] a dataset of 41 chemicals was re-classified in terms of chemical mechanistic classes. A larger set of 106 chemicals has been updated to extend these ideas in [16]. Work is underway to explore this further on an extended dataset of 210 chemicals published in [18]. The LLNA involves topical application of the test chemical to mouse ear skin followed by quantitative measurement of the T-cell proliferation response in the draining lymph node, assessed as a function of the incorporation of tritiated thymidine. The method is described in detail by Basketter et al. [19-21].

More work is on-going and still needs to be carried out in this field but it remains useful to highlight that much is already understood about the common reactions that are likely drive sensitisation and that these types of rules could be potentially encoded into a software program for more routine use. Examples of specific reaction schema have already been proposed in [17,22] (see Scheme 1).

Scheme 1. Reaction mechanistic applicability domains



Some available expert systems such as Derek for Windows (DfW) (LHASA Ltd, Leeds, UK) have successfully encoded common reactions, albeit to a lower level of granularity. DfW is a tool that describes structure-toxicity relationships for a number of endpoints though skin sensitisation is one of the best characterised. The skin sensitisation knowledge base in Derek was initially developed in collaboration with Unilever in 1993 using its historical database of GPMT data for 294 chemicals and contained approximately

forty alerts [23]. The current Derek knowledge base (version 9) has sixty-four alerts for skin sensitisation [24].

Structural chemical reaction schema and Derek rules are just two illustrations where human expert knowledge about skin sensitisation mechanisms has been successfully described.

3.1.2. Carcinogenicity

In the area of carcinogenicity, much work has been undertaken to study the different mechanisms of action. Carcinogens can be grouped into two main categories – genotoxic and non genotoxic (epigenetic carcinogens). The genotoxic carcinogens are characterised by covalent binding to DNA. Genotoxic mutagens also fall into this category as mutation is often one of the first steps in the development of cancer. The epigenetic carcinogens do not cause DNA damage directly, are usually negative in standard mutagenicity assays and act through a range of different mechanisms.

Genotoxic mutagens and carcinogens act as electrophiles and can be characterised by many of the standard chemical reactions as have been described for skin sensitisation previously. A number of human expert rules have been derived in the form of structural alerts (SAs). Well known examples include carbonium ions (alkyl-, aryl-, benzylic-), nitrenium ions, epoxides and oxonium ions, aldehydes, polarised α,β-unsaturated fragments, peroxides, free radicals, and acylating intermediates [25,26]. Examples of structural alerts for mutagenicity have been combined in Ashby's supermutagen model [25]. The supermutagen model is useful to identify potential carcinogens, though it should be noted that it is not an exhaustive list of all possible structural alerts.

As well as SAs, other factors identified as key determining factors in the expression of carcinogenicity include a) Molecular Weight (MW): larger chemicals are less likely to partition significantly; b) physical state, which can have an impact on the ability of a chemical reach the target site; c) solubility: highly hydrophilic compounds are generally poorly absorbed and, if absorbed, readily excreted; d) chemical reactivity: compounds which are "too reactive" may not be carcinogenic due to spontaneous hydrolysis, polymerisation or reaction with other cellular constituents; e) geometry, planar compounds, with an electrophilic functional group and favourable size, are able to intercalate into DNA. Examples of potent carcinogens include polycylic aromatic hydrocarbons; f) metabolism: metabolism can activate and detoxify chemical carcinogens: Knowledge of the metabolic pathway of a chemical can substantially enhance the accuracy of structure-activity analyses [27].

Expert systems such as Derek and Oncologic have attempted to encode a lot of SA information and human expert rules together. For example Derek contains many of the rules as defined by Ashby Tennant as well as the US Food & Drug Administration (US-FDA) structural alerts. In Derek (Version 9) there are 82 structural

alerts for mutagenicity and 53 alerts for carcinogenicity. Oncologic® is an expert system that assesses the potential of chemicals to cause cancer. Developed in co-operation with the US EPA and marketed by Logic Chem Inc, Oncologic® predicts the potential carcinogenicity of chemicals by applying SAR analysis and incorporating what is known about the mechanisms of action and human epidemiological studies. Oncologic® has the ability to reveal its line of reasoning just as human experts can. After supplying the appropriate information about the structure of the compound, an assessment of the potential carcinogenicity and the scientific line of reasoning used to arrive at the assessment outcome are produced. This information provides a detailed justification of a chemical cancer causing potential. Oncologic® can evaluate the following classes of compounds, fibres, polymers, metals, metalloids, and metal containing compounds as well as organic chemicals.

Derek and Oncologic™ are just two of the expert systems available where knowledge about the behaviour of genotoxic carcinogens and mutagens has been encoded for routine use.

Another example of human expert rules are the Cramer decision rules [28] applied in the area of Threshold of Toxicological Concern (TTC).

3.2. TTC and Structural Based Approaches

The TTC is a concept that refers to the establishment of a level of exposure for all chemicals below which there would be no appreciable risk to human health, providing a possible basis for waiving testing based on knowledge of exposure limits. A threshold is based from a statistical analysis of the toxicological data of a broad range of different and/or structurally related chemicals and extrapolation of the no effects dose obtained from the underlying animal experiments for these chemicals considered to be of negligible risk to human health.

The TTC concept has been incorporated into some risk assessment processes in regulatory schemes but its use is perhaps not so extensive. Areas where it has been explored are food additives and food contact agents in addition to impurities in pharmaceuticals.

There exist two principal approaches: either a general TTC is developed based principally on carcinogenicity data or a TTC based on structural information compared with toxicological data of chemicals (“the decision tree approach”) is derived for non-carcinogenic endpoints. The structural information used is typically that of Cramer [28]. The procedure utilises recognised pathways for metabolic deactivation and activation, data on toxicity and the presence of a substance as a component of traditional foods and as an endogenous metabolite. The substances are classified in to one of 3 categories. Class 1 are substances of simple chemical structure with known metabolic pathways and innocuous end products which would suggest a low order of oral toxicity e.g. butyl alcohol. Class 2 contains substances that are intermediate that is they possess

structures that are less innocuous than in the Class 1 but they do not contain structural features that are suggestive of toxicity like those of Class 3. Members of Class 2 include compounds such as allyl propionate or methyl 2-octynoate thus they may contain reactive functional groups. Class 3 are substances of a chemical structure that permit no strong initial impression of safety and may even suggest a significant toxicity. Examples include benzoin or 2-phenyl-3-carbethoxyfuran.

Toxtree, developed by Ideaconsult Ltd under ECB contract, is a freely available application from the ECB website (http://ecb.jrc.it/QSAR) which is able to estimate different types of toxic hazard by applying structural rules. Toxtree includes options for applying the Cramer decision tree [28] and more recently the Verhaar scheme [29-31].

3.3 Analogue-based approaches

Another approach for formulating a chemical grouping is through a systematic search using available analogue searching tools. The starting point for this bottom-up approach is either for a ("seed") chemical or else a small set of ("seed") chemicals. There are a number of tools both freely and commercially available to facilitate this type of searching. A description of a few of these tools is provided below.

TOXNET (http://toxnet.nlm.nih.gov/) managed by the Toxicology and Environmental Health Information Program (TEHIP) in the Division of Specialized Information Services (SIS) of the National Library of Medicine (NLM), is a free web-based system of integrated databases on toxicology, hazardous chemicals, environmental health and related areas. Several databases can be queried through TOXNET: HSDB, IRIS, ITER, GENE-TOX, CCRIS, HazMap, Household Product Database, TOXMAP, TOXLINE, DART, TRI, ChemIDPlus. The last of these has the facility to perform structural, substructural and similarity based searches to retrieve analogues and their associated data. The database contains over 368,000 chemical records, of which 200,000 include chemical structures. ChemIDPlus is accessible from http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?CHEM. Whilst the (sub)structural searches are self evident, it is perhaps worth explaining what is meant by a "similarity" search. The most common similarity index implemented in the majority of available structure searchable databases is the Tanimoto index. This compares one chemical structure with another on the basis of fingerprints to arrive at a number between zero and one. An index of 0 signifies that the two chemicals have nothing in common whereas a value of 1 reveals them to be identical.

The Tanimoto index is defined as follows:
number of bits in common between A and B /(number of bits in molecule A + number of bits in molecule B - number of bits in common between A and B)

ChemFinder (www.chemfinder.com) is another free chemical searching tool that has been on-line since 1995. The index provides chemical structures, physical properties, and hyperlinks to other data sources such as RTECS, TOXNET. Searching can be done by (sub)structure using a free plug in as well as on the basis of Tanimoto similarity.

The **Danish (Q)SAR Database,** freely available from the European Chemicals Bureau website (http://ecb.jrc.it/QSAR), provides a facility for (sub)structural searching to retrieve analogues. The system provides no experimental test data but provides estimated values for a number of human health and environmental fate effects for chemicals in various regulatory inventories.

The **Analog Identification Methodology (AIM)** has been developed by the U.S. Environmental Protection Agency to facilitate read across and chemical grouping by identifying chemical analogues that have existing test data publicly available. AIM is a web-based, computerized tool that identifies chemical analogues based on structure. The tool also provides the user with pointers or links to publicly available experimental data on the closely related chemical(s). AIM identifies chemical analogues from a default database that currently contains 31,031 compounds that have some type of toxicity data publicly available. AIM employs a fragment-based search method to identify analogous compounds using a set of 645 pre-defined fragments and correction factors, and a "three-pass" searching strategy to locate structures through defined rules and allowable substitution patterns for different types of structural features. AIM can be searched on the basis of structure, SMILES or CAS number, though it cannot be searched by chemical name.

The tool provides a simple means of identifying analogues that have some kind of toxicity data available, but it does not categorise or rank the analogues returned. This approach leaves it to individual users need to determine when a specific analogue is suitable for a specific assessment, as the determination of what structure is 'appropriate' can vary depending on the endpoint assessed.

The available test data is accessed in the form of hyperlink pointers. The data is not structured in any way and cannot be downloaded into Excel or other tools for analyses. Some hyperlinks point to a general webpage, e.g. IUCLID homepage or RTECS homepage, so the user may need the appropriate licenses to be able to extract available information. Other links take the user directly to the data source. Thus, the pointer informs that there is a record for the chemical, but does not always indicate the specific type of data available.

AIM allows users to rapidly categorize multiple chemicals, focus available resources, facilitate read across, and streamline assessment exercises.

(Note: AIM is not yet available for public release, anticipated for release in 2007).

Some analogue search facilities are sold as add-ins within MS Office based programs such as Excel including Accord (http://www.accelrys.com/products/accord/) or Chemfinder for Excel. Whereas with systems such as ChemID, Chemfinder which are populated with chemical structures, tools such as Chemfinder for Excel, Chemfolder (ACD Labs) (https://www.acdlabs.com/products/chem_dsn_lab/chemfolder/features.html) are incumbent on the user having a starting database, inventory or dataset of structures on which to conduct a search and make a selection.

Other more sophisticated tools include **Leadscope®** which is a software tool developed and commercialised by Leadscope Inc (www.leadscope.com). It possesses a unique chemical hierarchy containing over 27,000 chemical fingerprints which represent functional groups, chemical groupings, pharmacophores. The software can be purchased with a toxicity database and/or known drugs database. The toxicity database contains integrated information on over 160,000 chemical structures from multiple sources including FDA PAFA Database, National Toxicology Program (NTP), RTECS®, DSSTox Carcinogenicity Potency Database (CPDB). The databases cover a range of endpoints including acute and multiple dose studies such as subchronic liver toxicity, carcinogenicity, genetic toxicity, reproductive toxicity and irritation. The database can be searched by structure (such as substructure or similarity), type of study, toxic effect, species, sex, dosage, duration and route of exposure. Results can be viewed and exported in tables such as Excel.

3.4 Unsupervised Approaches

Unsupervised approaches involve the use of statistical techniques to split a dataset of chemicals into smaller groups – i.e. a top-down approach. The approach relies on a starting dataset/inventory of chemicals and computing different numerical chemical parameters (such as geometrical, topological, structural, physicochemical, electronic descriptors) for those chemicals or characterising them through the use of fingerprints/structural features. No assumptions are made about which parameters are more or less relevant towards an endpoint or property thus a grouping can be performed on the basis of as much chemical information as possible or only parameters that are thought to be influential for a given (eco)toxicological endpoint.

For example a number of descriptors (say 100) could be calculated for a dataset/inventory of chemicals and a range of statistical approaches used to split the set into groups of "similar" chemicals where similarity is represented by similarity with respect to all 100 descriptors calculated. Alternatively expert judgement could be applied to select only those descriptors thought to be relevant to the endpoint of concern in order to make a grouping to identify those chemicals similar in toxicity based on the relevant descriptors selected. For example, skin sensitisation is dependent on chemical reactivity and to a limited extent partitioning. A number of (Q)SARs for skin sensitisation have been developed relating skin sensitisation potency to descriptors to

modelling reactivity and partitioning [22]. These descriptors could be used as a basis of grouping similar chemicals where similarity is represented by the likely sensitisation effect. Potentially useful descriptors might include Log P (logarithm of the octanol/water partition coefficient), Taft coefficients, LUMO (energy of the lowest unoccupied molecular orbital) to model partition and reactivity respectively [22].

The sorts of statistical approaches that can be used to split the dataset/inventory vary. Common techniques include principal components analysis (PCA), clustering and self organising maps (SOMs).

A short description of these is given below.

3.4.1. Principal Components Analysis (PCA)

PCA is a technique that aims to reduce the data dimensionality whilst retaining the information content. This is particularly useful in exploratory analysis enabling high dimensionality data to be readily visualised. Thus it becomes possible:

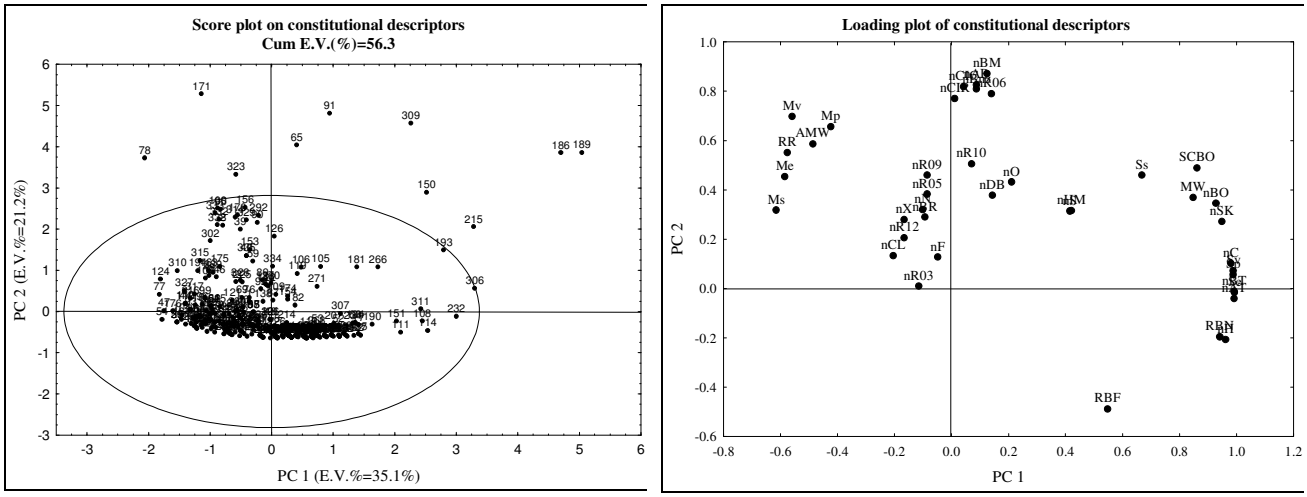
- to identify clusters of similar chemicals and whether these provide some insights into differing mechanisms of action
- to highlight chemicals that might be potential outliers
- to gain insights about which descriptors might be relevant
- to remove descriptors that do not provide useful information (so-called noisy variables)
- to gain an understanding of which descriptors are "related" to each other (co-linear)

The technique essentially comprises two main operations, a translation followed by a rotation, thus data for a set of descriptors (variables) are "projected" into a new coordinate system. The axes of this coordinate system are termed principal components (PC) which are linear combinations of the original variables. The main properties of these components are thus; the first principal component explains the maximum variance in the dataset with subsequent components describing the maximum part of the remaining variance subject to the condition that all principal components are orthogonal to one another. In addition as many principal components are extracted as the original starting set of variables. For example if 10 descriptors were variables in the original coordinate system, 10 PCs would be produced after a PCA. The first component (PC1) explains the maximum variance – for example 70%, hence the second PC 25%, the third 3% and so on until all (100%) of the variance is explained. In other words, the majority of information (95%) is captured in only 2 PCs instead of 10 descriptors. This makes visualisation in 2D easy and insights to be extracted much more efficiently. It also means that information is appropriately formatted (pre-processed) for other statistical techniques that are very sensitive to inter-correlating descriptors.

The output of a PCA comprises a matrix of scores and loadings. For example if a PCA was conducted on a set of chemicals then a scores plot would enable a visualisation of how the chemicals were clustered in 2 or 3 dimensions (see Figure 1a). The loadings plot would provide some insights in the complex nature of the PCs themselves, the directionality of the variables and how they were clustered together (see Figure 1b) [32].

Figure 1 - Score plot of PC1 *vs* PC2 calculated from constitutional descriptors (taken from M. Pavan – with permission)

Figure 1 - Loading plot of PC1 *vs* PC2 calculated from constitutional descriptors (taken from M. Pavan – with permission).



3.4.2. Clustering

Clustering is an extremely useful *unsupervised learning* technique, which aims to discover *structures* in a collection of data without explaining why they exist. A loose definition of clustering could be "the process of organising objects (chemicals) into groups whose members are similar in some way". A *cluster* is therefore a collection of objects (chemicals) which are "similar" between themselves and are "dissimilar" to the objects belonging to other clusters. Clustering algorithms may be classified as Exclusive Clustering (1), Overlapping Clustering (2), Hierarchical Clustering (3) and Probabilistic Clustering (4).

In (1), data are grouped in an exclusive way, so that if a certain data point belongs to a definite cluster then it can not be included in another cluster. In (2), the overlapping clustering, uses fuzzy sets to cluster data, so that each point may belong to two or more clusters with different degrees of membership. In this case, data will be associated to an appropriate membership value. A hierarchical clustering algorithm is based on the union between the two nearest clusters. The starting condition is realised by setting every data point as a cluster. After several iterations final clusters are realised. (4) relies on a completely probabilistic approach. Here we will only highlight two of the most common clustering techniques, K-means (an exclusive clustering technique) and hierarchical clustering.

K-means proceeds by performing an iterative alternating fitting process to form the specified number clusters. It first selects a set of *n* points (the cluster seeds or chemicals) as an initial guess of the means of the clusters. Each observation is assigned to the nearest seed to form a set of temporary clusters. The seeds are then replaced by the cluster means, and the process continues until no further changes occur in the clusters.

K-means method is intended for use with larger datasets, up to approximately 100,000 observations (or in this context 100,000 chemicals) [33].

Hierarchical clustering by agglomerative nesting

Hierarchical clustering provides a more specific control by assigning every single compound to a cluster of compounds. It does not require any prior assumption about the number of clusters and after the process is completed, the compound closest to the centre of the cluster is selected as a representative compound [33].

Agglomerative nesting starts with each compound in a separate cluster and proceeds by merging the closest pair of clusters at each iteration. Some examples (employing different distance measures) are highlighted below.

1. Average linkage

2. Single linkage (nearest neighbour)

3. Complete linkage (furthest neighbour)

4. Ward's minimum variance method

Average Linkage. A *pair* consists of two observations, *one from each cluster*. The distance between two clusters is defined to be the average distance between all pairs in the two clusters. This method tends to join clusters with small variances and is slightly biased toward producing clusters with the same variance. The average linkage method is one of the better methods for structure clustering.

Complete Linkage (Furthest Neighbour). The distance between two clusters is defined to be the maximum distance between an observation in one cluster and an observation in the other cluster. This method is biased toward producing clusters that divide the space approximately equally and can be distorted by moderate outliers.

Single Linkage (Nearest Neighbour). The distance between two clusters is defined to be the minimum distance between an observation in one cluster and an observation in the other cluster. This method has the ability to detect elongated and irregular clusters.

Ward's Minimum Variance Method.

The distance between two clusters is defined to be the ANOVA sum of squares between the two clusters. In each step, the within-cluster sum of squares is minimized by merging two clusters from the previous step. Ward's method tends to join cluster with a small number of observations and is strongly biased toward producing clusters with roughly the same number of observations. This method is very sensitive to outliers [33].

3.4.3. Unsupervised Self Organising Maps (SOMs) or Kohonen networks

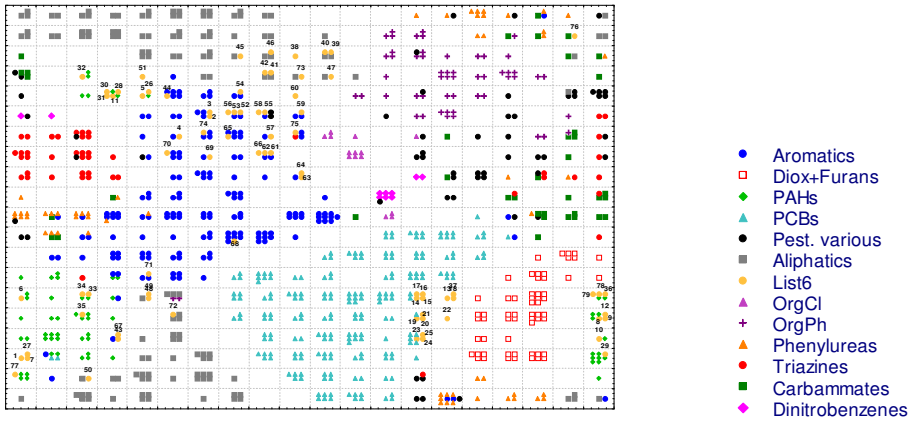
A SOM is a learning algorithm which represents high-dimensional data in a low-dimensional form without losing any of the 'essence' of the data. The data is organised on the basis of similarity by putting entities geometrically close to each other. The main goal of the neural network is to map compounds from n-dimensional into 2-dimensional space. Compounds similar in the original multidimensional space are close to each other in the map. A SOM network has only two layers: the input layer, and an output layer known as the topological map (top-map) layer. The units in the topological map layer are usually laid out in two dimensions. SOM networks are trained using an iterative algorithm. Starting with an initially-random top map, the algorithm gradually adjusts this to reflect the clustering of the training data. The iterative training procedure also arranges the network so that units representing centres close together in the input space are also situated close together on the topological map. The basic algorithm runs through a number of epochs (time when the whole training set is presented to the network), each time and for each case, the following occurs:

- a winning neuron is selected (the one who's centre is nearest to the input case)
- the winning neuron is adjusted to be more like the input case (a weighted sum of the old neuron centre and the training case)

The topological ordering property is achieved by adding the concept of a neighbourhood to the algorithm. The neighbourhood is a set of neurons surrounding the winning neuron. The effect of this neighbourhood is that initially quite large areas of the network are "pulled towards" training cases quite substantially. The network develops a crude topological ordering, with similar cases activating clumps of neurons in the topological map. As epochs pass, finer distinctions within areas of the map can be drawn.

Figure 2 shows the output of a K-NN for a dataset of 974 chemicals described by 10 variables (PCs calculated on 839 molecular descriptors) and coloured coded by chemical functionality.

Figure 2 – Example TOP map (taken from M. Pavan– with permission)



3.5. Supervised learning approaches

Supervised learning approaches are another example of a top-down approach. There is some degree of overlap in the statistical techniques that might be applied in supervised learning. The key difference is that information about the activity/toxicity of chemicals is taken into account in addition to the structural/descriptor information

For example clustering techniques may still be employed but the criterion is that the clusters are extracted to discriminate for the toxicity present.
Other techniques might include recursive partitioning where the aim will be to find active or statistically correlated subsets based on the presence or absence of a particular combination of substructural features/fingerprints. It has been successfully used for molecular diversity and similarity analysis. Other techniques useful for formation of grouping incorporating activity profiles include ranking approaches. A brief overview of both is provided below.

3.5.1. Recursive partitioning

Recursive partitioning (RP) is a technique that works by trying to find a single predictor whose different values split the compounds up into subgroups of more homogeneous activity. These groups are then analyzed as if they were an original data set and are divided in subgroups (hence the recursive part in RP). The primary outcome of a recursive partitioning analysis is a dendrogram where each box or 'node' represents a group of chemicals [33].

3.5.2. Ranking approaches

Ranking approaches are multivariate sorting techniques capable of: a) prioritising/ordering a set of chemicals; and b) grouping a set of chemicals according to more than one variable.

There are two main types of ranking technique, total order and partial order ranking. The total ordering ranking fully orders a dataset regardless of whether there are disagreements between the variables. For some total ordering approaches, a function for each endpoint variable is used to transform it onto the same scale. Then each endpoint variable is independently transformed into a desirability (for the desirability function approach) or utility (for the utility function approach) by an arbitrary function which transforms the actual value for a chemical between 0 and 1. A global desirability/utility is then derived by combining the individual desirabilities/utilities by a geometrical mean (desirability function), or arithmetic mean (utility function). The choice of function will depend on the specific context - whether differences between the

endpoint variables wish to be emphasised or whether greater discrimination between groups is desired. An alternative function is the dominance function approach. This approach is based on the comparison of the state of the different endpoint variables for each pair of chemicals. The approach does not require a transformation instead whether the best condition is satisfied by a minimum or maximum value of the selected endpoint variable needs to be determined. Figure 3a shows an example of a total ordering ranking using a desirability function. (Figure used with permission from M. Pavan)

However whilst total order ranking methods are able to force a total ranking, they lose information on conflicting criteria during the process. An alternative technique is the partial order ranking method which recognises that not all chemicals can be directly compared with others, due to the presence of conflicting criteria. The Hasse diagram technique has proved a useful tool to perform partial order ranking (POR). It was introduced into the environmental sciences field by Halfon [34-36] and has subsequently been refined by Brüggemann [37-38]. Each chemical is represented by a small circle. Comparable chemicals which belong to an ordered relation are linked, while incomparable chemicals are not connected. An example is shown in Figure 3b of such a Hasse diagram. More information on ranking approaches in chemical sciences can be found in [39].

Figure 3b: Example of a total order ranking using the desirability function (used with permission from M.Pavan)

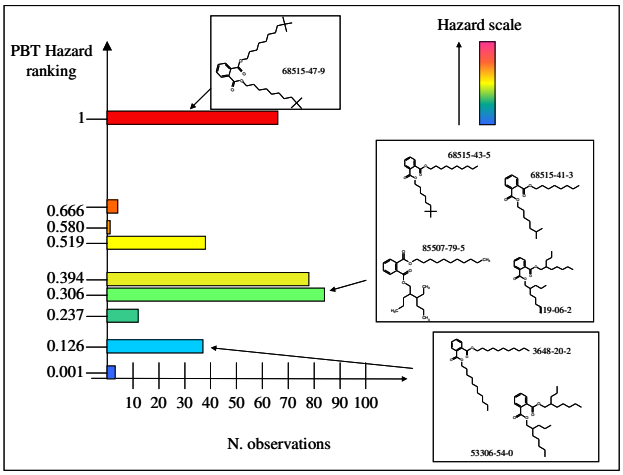
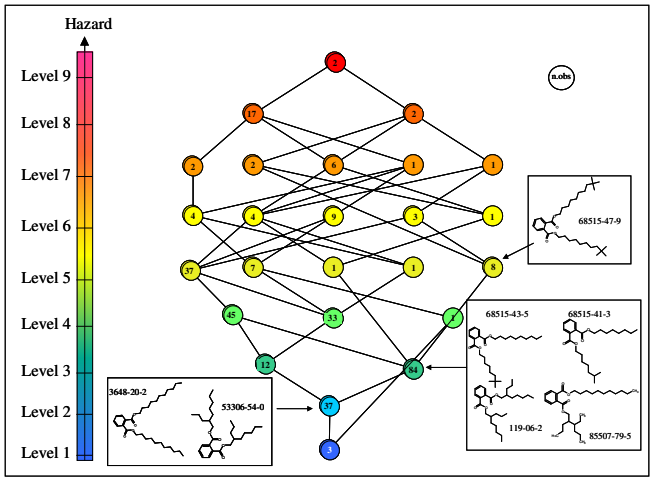


Figure 3b: Example of a Hasse diagram (used with permission from M.Pavan)



4. Main Conclusions - Flowchart

As has already been described, there are a number of grouping approaches available – knowledge-based, analogue-based, unsupervised, and supervised. The purpose of the grouping and starting point of information will determine which approaches are most appropriate to use. In Table 1 we have attempted to outline some

of the potential situations and hence methods. It is worth noting that no single solution or recommendation is made here. Most likely, a battery of different approaches used iteratively will provide the most pragmatic and practical means of forming chemicals groupings. The flowchart below (in Scheme 2) attempts to capture the possible steps as part of a workflow but this should be regarded as flexible and not the only possible approach. The first step considers the available information – both the chemical(s) interest and the context i.e. the endpoint. If the available starting information was a small set of chemicals or a single chemical – a build up of information to identify other similar chemicals would be a pragmatic next step. This could be carried out by conducting an analogue search, or by applying toxicological insight about a chemical class. This type of approach is termed here as a "bottom up" approach since it starts from a small amount of information which is then extended and increased to result in an initial grouping. If the available starting information was a large inventory of chemicals or a large dataset, the next step would be a rationalisation to form smaller manageable groups. This could be done using a number of different techniques such as (un)supervised statistical methods. This is termed here as "top down" in that a large amount of information is dissected to formulate a number of initial groupings.

Irrespective of these starting points, the initial groupings are then evaluated and reviewed. (Q)SAR approaches can be helpful in this process to refine the groupings formed. This part may be iterative.

Scheme 2: Proposed Category Flowchart

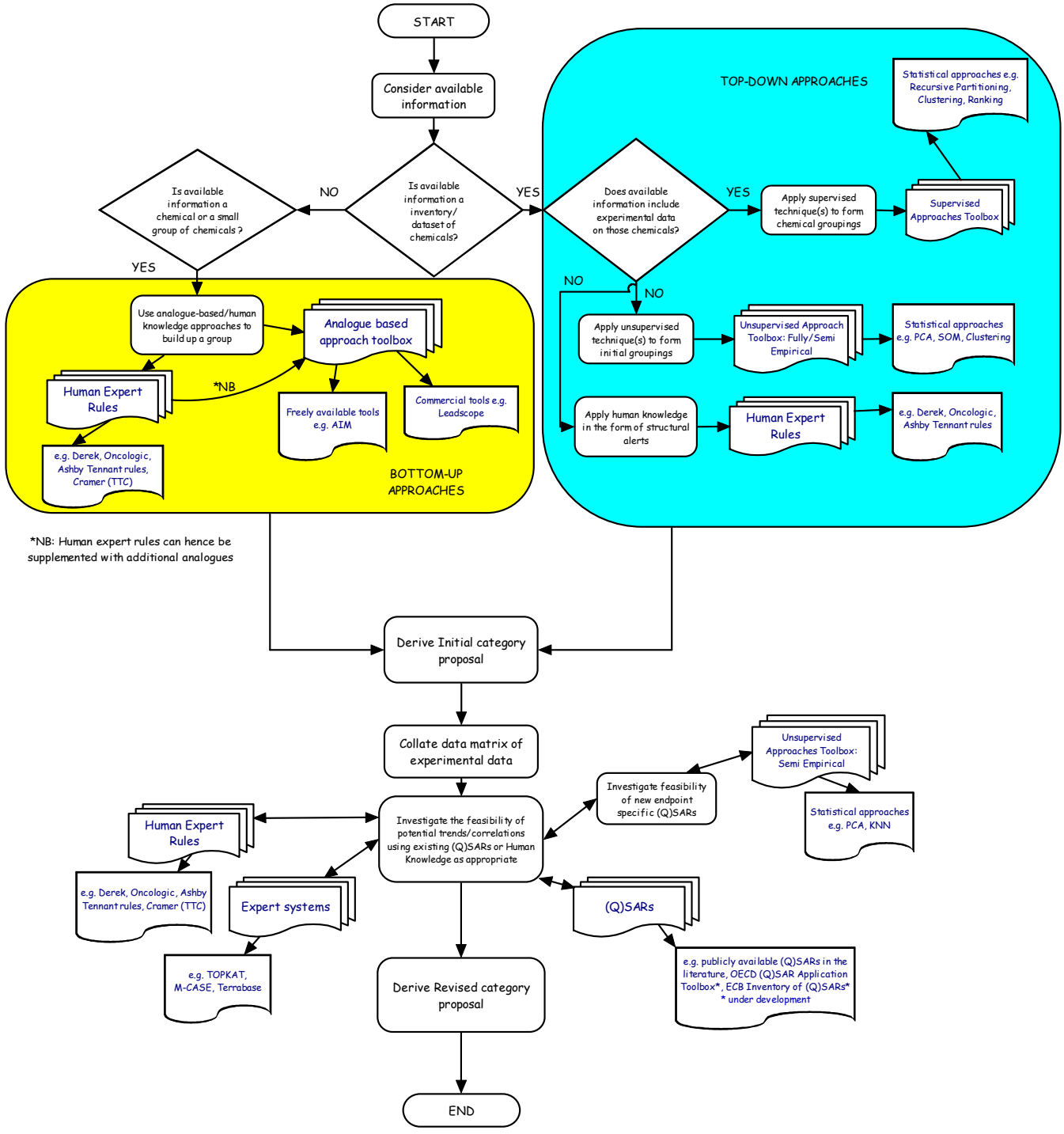


Table 1. Table of Potential Approaches

Type of Approach	Application	Advantages	Conditions of Use
Knowledge-based approaches: Human Expert Rules	• Top-Down: Rules can be used as "seeds" to formulate initial groupings for a starting inventory of compounds (without an endpoint context) • Bottom-up: Rules can be used to form endpoint specific groupings which can hence be supplemented by analogue based search	• Pragmatic means of performing structural grouping • Knowledge of structure toxicity relationships can be expressed in terms of rules with clearly defined applicability domains	• Relies on knowledge being sufficiently well developed to be encoded into meaningful rules
Analogue approaches (Bottom-Up)	• Analogue search tools are available free or commercial to facilitate systematic searching of analogues	• Can be used to confirm human expert rules	• Tend to retrieve structurally similar analogues
Unsupervised Approaches (Top-Down)	• Enables a pragmatic and empirical grouping to be performed for a set of chemicals characterised by a large set of different descriptors or fingerprints • A semi-empirical grouping can be performed where endpoint	• Fast systematic means of performing groupings in the absence of available experimental data.	• Only information about the chemical is used in the analysis • Different statistical techniques such as PCA, Clustering, Neural Networks can be used to formulate the groupings • Relies on a starting inventory of

	relevant descriptors are used to account for likely similar behaviour.	chemicals as well as associated calculated properties or fingerprints. • Results are very dependent on starting descriptors and what information is captured within those descriptors
Supervised approaches (Top-Down)	• Semi-empirical technique using a combination of chemical information together with experimental data. • Statistical tools are used to help formulate the groupings.	• Relies on available experimental data and chemical information, a starting inventory of chemicals and database is required.

5. Recommendations for further work

The flowchart presented is theoretical one based on the discussions and thinking held at the time of the workshop. The flowchart merely serves to capture and summarise the discussions in a visual form. It is not a prescriptive proposal nor has it been tried in practice with examples. It will need to undergo refinements as experience is gained through the formation of categories either under REACH or from in-house investigative research performed using different QSAR approaches. Further work will involve identifying appropriate scenarios where a range of QSAR tools and approaches can be used to evaluate the flowchart – testing out the different starting points and evaluating the sorts of initial groupings formed.

Some aspects of the flowchart are better supported by existing computational tools than others. Analogue searching is one area that would benefit from additional functionality. Whilst there are a number of tools available that encode different similarity measures, there is no tool that enables a comprehensive set of indices to be generated and applied for the retrieval of biologically similar analogues. To that end, the ECB has commissioned the development of a software tool (Toxmatch) that will encode different types of similarity measures and will provide guidance on how they can be applied to specific endpoints as examples (endpoints such as skin irritation and aquatic toxicity will be included to demonstrate capability). This prototype tool should be available in 2007 and is one of the key follow up activities to this workshop.

6. References

1. European Commission 2003. Proposal for a Regulation of the European Parliament and of the Council concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency and amending Directive 1999/45/EC and Regulation (EC) {on Persistent Organic Pollutants}. Website: http://europa.eu.int/comm/enterprise/chemicals/chempol/whitepaper/reach.htm
2. Council of the European Union. (2006). Common position adopted by the Council with a view to the adoption of a Regulation of the European Parliament and of the Council concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC of the European Parliament and of the Councol and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. Text agreed by the Council of Ministers on 12 June 2006. http://europa.eu.int/comm/enterprise/reach/index_en.htm
3. Organisation for Economic Cooperation and Development (2004). OECD Principles for the Validation, for Regulatory Purposes, of (Quantitative) Structure-Activity Relationship Models. [http://www.oecd.org/document/23/0,2340,en_2649_34379_33957015_1_1_1_1,00.html]
4. Worth AP, Bassan A, Gallegos A, Netzeva TI, Patlewicz G, Pavan M, Tsakovska I, Vracko M. 2005. The Characterisation of (Quantitative) Structure-Activity Relationships: Preliminary Guidance. ECB Report EUR 21866 EN, 95pp. Ispra, Italy: European Commission, Joint Research Centre.
5. OECD 2005. Section 3.2 of the OECD Manual for Investigation of HPV Chemicals: Chapter of guidance document on the formation and use of chemical categories. http://www.oecd.org/document/7/0,2340,en_2649_34379_1947463_1_1_1_1,00.html
6. Anon 2005. Report of the SPORT (Strategic Partnership on REACH Testing) pilot project: “The SPORT Report: Making REACH work in practice”. Website: http://www.sport-project.info
7. Dearman RJ, Kimber I. 2003. Factors influencing the induction phase of skin sensitization. American Journal of Contact Dermatitis 14, 188-194.
8. Kimber I, Dearman RJ. 2003. What makes a chemical an allergen? Annals of Allergy, Asthma and Immunology 90, 28-31.
9. Smith Pease CK, Basketter DA, Patlewicz GY. 2003. Contact allergy: the role of skin chemistry and metabolism. Clinical and Experimental Dermatology 28 (2), 177-83.
10. Pease Smith CK. 2003. From xenobiotic chemistry and metabolism to better prediction and risk assessment of skin allergy. Toxicology 192: 1-22.
11. Landsteiner K, Jacobs JL. 1936. Studies on the sensitisation of animals with simple chemicals III. Journal of Experimental Medicine 64: 625-639.
12. Roberts DW, Williams D L. 1982. The Derivation of Quantitative Correlations between Skin Sensitisation and Physico–Chemical Parameters for Alkylating Agents and their Application to Experimental Data for Sultones. Journal of Theoretical Biology 99: 807 – 25.

13. Patlewicz G, Roberts DW, Walker JD. 2003. QSARs for Skin Sensitisation Potential of Aldehydes and Related compounds. Quantitative Structure-Activity Relationships & Combinatorial Science 22: 196-203
14. Aptula N, Patlewicz GY, Roberts DW. 2005. Skin Sensitisation: Reaction Mechanistic Applicability Domains for Structure-Activity Relationships. Chemical Research in Toxicology 18(9):1420-1426
15. Roberts DW, Aptula N, Patlewicz GY. 2006. Mechanistic Applicability Domains for Non-Animal Based Toxicological Endpoints. QSAR Analysis of the Schiff Base Applicability Domain for Skin Sensitization. Chemical Research in Toxicology 19(9): 1228-1233
16. Roberts DW, Aptula N, Patlewicz GY. 2006. Structure-Activity Relationships in Skin Sensitisation: Reaction Mechanistic Applicability Domain Classification for a Published Data Set of 106 Chemicals tested in the Mouse Local Lymph Node Assay. Chemical Research in Toxicology, *accepted for publication*
17. Aptula N, Roberts DW, Patlewicz GY, Schultz TW. 2006. Non-Enzymatic Glutathione Reactivity and In Vitro Toxicity: A Non-Animal Approach to Skin Sensitization. Toxicology in Vitro 20(2):239-47
18. Gerberick GF, Ryan CA, Kern PS, Schlatter H, Dearman RJ, Kimber I, Patlewicz G, Basketter DA. 2005. Compilation of historical local lymph node assay data for the evaluation of skin sensitization alternatives. Dermatitis 16(4):157-202
19. Kimber I, Basketter DA. 1992. The murine local lymph node assay; collaborative studies and new directions: A commentary. Food Chemical and Toxicology 30: 165-169.
20. Basketter DA, Blaikie L, Dearman RJ, Kimber I, Ryan CA, Gerberick GF, Harvey P, Evans P, White IR, Rycroft RJG. 2000. Use of the local lymph node assay for the estimation of relative contact allergenic potency. Contact Dermatitis 42: 344-348.
21. Basketter DA, Roberts DW, Cronin MTD, Scholes EW. 1992. The value of the LLNA in quantitative structure activity investigations. Contact Dermatitis 27: 137-142
22. Roberts DW, Aptula N. 2006. Mechanistic Applicability Domains for Non-Animal Based Prediction of Toxicological Endpoints. General Principles, and Application to Reactive Toxicity. Chemical Research in Toxicology, 19(8): 1097-105
23. Barratt MD, Basketter DA, Chamberlain M, Admans GD, Langowski JJ. 1994. An expert system rulebase for identifying contact allergens. Toxicology in Vitro. 8: 1053-1060.
24. Barratt MD, Langowski JJ. 1999. Validation and subsequent development of the Derek skin sensitization rulebase by analysis of the BgVV list of contact allergens. Journal of Chemical Information & Computer Science 39: 294-298.
25. Ashby J. 1985. Fundamental structural alerts to potential carcinogenicity or noncarcinogenicity. Environmental Mutagenicity 7: 919-921.
26. Woo YT, Lai D, McLain JL, Manibusan MK, Dellarco V. 2002. Use of mechanism-based structure-activity relationships analysis in carcinogenic potential ranking for drinking water disinfection by-products. Environmental Health Perspectives 110 Suppl 1:75-87.
27. Benigni R. 2005. Structure-activity relationship studies of chemical mutagens and carcinogens: mechanistic investigations and prediction approaches. Chemical Reviews 105: 1767-1800.
28. Cramer GM, Ford RA, Hall RA. 1978. Estimation of toxic hazard – a decision tree approach. Food and Chemical Toxicology 16: 255-276.

29. Verhaar HJM, van Leeuwen CJ, Hermens JLM. 1992. Classifying environmental pollutants. 1. Structure-activity relationships for prediction of aquatic toxicity. *Chemosphere* 25, 471-491.
30. Verhaar HJM, Mulder W, Hermens JLM. 1995. QSARs for ecotoxicity. In: Overview of structure-activity relationships for environmental endpoints, Part 1: General outline and procedure. Hermens JLM (Ed), Report prepared within the framework of the project "QSAR for Prediction of Fate and Effects of Chemicals in the Environment", an international project of the Environment; Technologies RTD Programme (DGXII/D-1) of the Europenan Commission under contract number EV5V-CT92-0211.
31. Verhaar HJM, Solbe J, Speksnijder J, van Leeuwen CJ, Hermens JLM. 2000. Classifying environmental pollutants: Part 3. External validation of the classification system. *Chemosphere* 40, 875-883.
32. Livingstone D. 1995. Data Analysis for Chemists. Oxford Science Publications.
33. Enterprise User Manual Version 2.2. 2001, 2002 Leadscope Inc.
34. Halfon E, Reggiani MG. 1986. On Ranking Chemicals for Environmental Hazard. Environmental Science & Technology 20: 1173-1179.
35. Halfon, E. (1989). Comparison of an Index Function and a Vectorial Approach Method for Ranking of Waste Disposal Sites. Environmental Science & Technology 23: 600-609.
36. Halfon E, Brüggemann R. 1998. On Ranking Chemicals for Environmental Hazard. Comparison of methodologies Proceedings of the Workshop on Order Theoretical Tools in Environmental Sciences 11-48.
37. Brüggemann R, Bartel H-G. 1999. A Theoretical Concept to Rank Environmentally Significant Chemicals. Journal of Chemical Information and Computer Sciences 39: 211-217.
38. Brüggemann R, Halfon E, Welz G, Voigt K, Steinberg CEW. 2001. Applying the Concept of Partially Ordered Sets on the Ranking of Near-Shore Sediments by a Battery of Tests. Journal of Chemical Information and Computer Sciences 41: 918-925.
39. Pavan, M. 2003. Total and Partial Ranking Methods in Chemical Sciences. PhD Thesis. University of Milano-Bicocca, Milan, Italy.

Appendix 1: Agenda

Day 1 – Chemical Similarity
(7 November)

09.00	**Start of the meeting Day 1**
09.00-09.30	Introduction of the participants. Aims and organisation of the meeting (Andrew Worth)
09.30-10.15	Chemical Similarity - an overview (Nina Jeliazkova)
10.15-10.45	Insights on Chemical Quantum Molecular Similarity Indices (Ana Gallegos)
10.45-11.15	**Coffee Break**
11.15-11.45	Chemical similarity in database searching (Val Gillet)
11.45-12.15	The concept of chemical categories (Brigitte Simon-Hettich)
12.15-12.45	Experiences in chemical series definition and chemical similarity (Aldo Benigni)
12.45-14.00	**Lunch**
14.00-14.30	From classification schemes for chemical structures to virtual biological profiling of chemical libraries (Jordi Mestres)
14.30-15.00	Introduction to the brainstorming and formulation of open questions (Grace Patlewicz)
15.00-15.30	**Coffee Break**
15.30-17.00	Discussion/ brainstorming on applicability of the techniques
17.00-17.30	Conclusions and recommendations.
17.45	**End of Day 1**

Day 2 – TTC
(8 November)

9.00	**Start of Day 2**
9.00-9.30	Review of Day 1
9.15-10.00	TTC - an OFAS perspective (Andrew McDougal (recorded presentation))
10.00-10.45	The Threshold of Toxicological Concern concept (Ian Munro)
10.45-11.15	**Coffee Break**
11.15-11.45	TTC - Literature review and applicability (Maria Wallén)
11.45-12.15	TTC - a SEAC perspective (Bob Safford)
12.15-12.45	TTC - Cramer classification scheme : a toolbox (Nina Jeliazkova)
12.45-14.15	**Lunch**
14.15-15.00	Overview of grouping (Chihae Yang)
15.00-15.30	Introduction to the brainstorming (Grace Patlewicz)
15.30-16.00	**Coffee Break**
16.00-16.30	Discussion/ brainstorming on applicability of the techniques
16.30-17.00	Conclusions and Recommendations – Report writing and next steps
17.00	**End of Day 2 and of the meeting – Transport to the airport**

Appendix 2: List of participants

Invited Experts

Name	Organisation	Country
Aynur (Nora) Aptula	Safety & Environmental Assurance Centre (SEAC), Unilever R&D	UK
Romualdo Benigni	Experimental and Computational Carcinogenesis Unit, Environment and Health Department, Dipartimento di Ambiente e connessa Prevenzione Primaria, Istituto Superiore di Sanità	Italy
Cecilia Bossa	Experimental and Computational Carcinogenesis Unit, Environment and Health Department, Dipartimento di Ambiente e connessa Prevenzione Primaria, Istituto Superiore di Sanità	Italy
Andrew McDougal	Division of Food Contact Notifications, Office of Food Additive Safety, FDA	USA
Agneta Falk-Filipsson	Swedish Chemical Inspectorate	Sweden
Val Gillet	Computational Informatics Research Group, Department of Information Studies, University of Sheffield	UK
Nina Jeliazkova	IDEA Consult Ltd.	Bulgaria
Jordi Mestres	Research Unit on Biomedical Informatics, Municipal Institute of Medical Research (IMIM)	Spain
Ian Munro	CANTOX Health Sciences International	Canada
Bob Safford	Safety & Environmental Assurance Centre (SEAC), Unilever R&D	UK
Brigitte Simon-Hettich	Institute of Toxicology, Merck KGaA	Germany
Maria Wallén	Swedish Chemical Inspectorate	Sweden
Chihae Yang	Leadscope® Inc.	USA

ECB Participants

Grace Patlewicz (Chairperson)	QSAR Group	Italy
Ana Gallegos Saliner	QSAR Group	Italy
Tatiana Netzeva	QSAR Group	Italy
Manuela Pavan	QSAR Group	Italy
Juan Riego Sintes	Testing Methods	Italy
Ivanka Tsakovska	QSAR Group	Italy
Andrew Worth	QSAR Group	Italy

Appendix 3: Summary of the ECB workshop

The meeting was chaired by **Grace Patlewicz** (ECB), who opened the workshop by welcoming the participants through a roundtable of introductions.

Presentations were made by several of the participants in order to provide an overview of ongoing activities from the perspective of different organisations (academia, industry, and regulatory organisations). This gave a perspective of some of the approaches available in the field of Chemical Similarity and TTC and how they were being applied. The presentations helped to structure the afternoon plenary discussions aimed at capturing potential strategies for Chemical Category development as well as research needs or opportunities.

Day 1 – Chemical Similarity

Andrew Worth (ECB) presented the aims, organisation and structure of the meeting. He briefly outlined the structure of the European Commission, the role of the JRC and that of ECB within the scientific and technical preparations for REACH. He presented the scope of the meeting namely, a review of approaches for chemical similarity and thresholds of toxicological concern. These approaches are of specific interest to the QSAR group since chemical similarity techniques could be potentially used to help classify chemicals into similarity-based chemical categories for read-across; and Threshold of Toxicological Concern for human health endpoints could help to evolve integrated testing strategies. He explained that the two topics (TTC and chemical similarity) had been combined into a single meeting, because they are basically both grouping approaches. Chemical similarity approaches provide a means of grouping chemicals for hazard identification (classification) purposes, whereas TTC approaches could be adapted to group chemicals according to their potency, i.e. provide a means of quantitative read-across.

Nina Jeliazkova (IDEA Consult Ltd.) presented a literature-based review on chemical similarity. She began by presenting similarity as an intuitive concept widely used in philosophy as well as many other disciplines. A meaningful, unambiguous and useful measure of similarity is needed to capture the resemblance in relation to the aspect to be described. She highlighted a myriad of different approaches for measuring the similarity between chemicals, from simple fingerprint counts, to 3D similarity including quantum chemistry field-based approaches. She stressed some of the main advantages and disadvantages of these different methods, depending on the numerical representation chosen for the molecular structures and the different types of similarity indices that are available. She concluded by highlighting several caveats for chemical similarity, in particular, how there is always a loss of information associated with any similarity measure; how some measures may not correctly represent the intuitive similarity between two chemicals; or even that structure may not be the sole factor for biological activity and that structurally similar molecules may still have differing mechanisms of action.

Ana Gallegos (ECB) presented some theoretical insights on the formulation of molecular similarity indices based on quantum mechanics calculations. She started by presenting the foundation of quantum similarity theory based on the characterisation of molecular structures by electronic density functions. She illustrated several approaches used to calculate first-order electronic density functions which minimise computational costs but preserve accuracy. The

atomic shell approximation (ASA), and the promolecular ASA (PASA) are examples of these. She also presented different algorithms for molecular superposition, based on the maximal similarity alignment rule or the topo-geometrical superposition rule. She introduced topological quantum similarity measures based on the classical topological representation of molecular structures by molecular graphs. She stressed the novelty of this approach in that by substituting classical topological two-dimensional matrices with quantum derived matrices, important three-dimensional information can be accounted for.

Val Gillet (University of Sheffield) presented chemical similarity techniques used in database searching and applied in the pharmaceutical industry. These measures are based on the calculation of the pairwise similarity between a known active molecule and each database compound, and the subsequent ranking of the compounds according to their similarity to the known active. She presented similarity measures based on the representation of compounds by two-dimensional fingerprints (vectors with the binary values of 0 and 1, accounting for the absence or presence of certain fragments), and using the Tanimoto index as a quantitative measure of similarity. She also presented a novel method based on the assignment of four properties to each functional group, encoded by triplets of strings, and the use of reduced graphs. She finally illustrated the theoretical basis with several virtual screening, and data fusion experiments, based on the combination of different rankings on the same sets of molecules.

Brigitte Simon-Hettich (Merck Institute of Toxicology) provided an overview of the chemical category concept from a toxicological point of view, including some examples from the notification of new chemicals in the EU. She introduced the chemical category concept based on its use within the US EPA and the OECD. The main advantages of categories are their potential savings in cost, time, resources, and animal experimentation. She illustrated the principles of the US EPA approach and the OECD approach with some examples. The OECD approach groups compounds which show a predictable pattern in physicochemical properties, environmental fate, environmental effects or human health effects in order to identify and fill in data gaps for relevant endpoints. She raised some questions and concerns related to categories based on common functional groups, metabolic pathways, and incremental changes in groups. For example, the practicality and utility of forming categories based on metabolic pathways was questioned. She also highlighted the need for chemical categories based on common mechanisms of action.

Romualdo Benigni (Istituto Superiore di Sanità) provided theoretical insights and practical examples based on the definition of chemical series, and on the use of chemical similarity in the context of categorizing biologically active chemicals. He started by raising the issue of why there is a need to define a valid chemical similarity measure and gradual scales of it. He highlighted the need for a subdivision between predictions of the biological activity of untested compounds from known QSAR into predictions within the spanned substituent space (SSS) and predictions outside the SSS. Presenting a series of real life analyses, he showed that different representations of the molecules can be highly correlated at the level of the universe of chemicals, while behaving very differently at the fine grain scale. This implies that the selection of the chemical descriptors has dramatic consequences on the issue of categorizing the chemicals. In addition, he showed that categories based only on the chemical theory do not encompass the toxicological properties. He concluded that the chemical similarity approaches for the categorization

of the chemicals should incorporate, and give a major weight to the mechanistic knowledge on the biological activity of the molecules.

Jordi Mestres (Municipal Institute of Medical Research) presented new challenges and achievements in the field of chemogenomics. He started by explaining the transition from mapping chemical and biological entities to obtain QSAR to using high throughput mapping techniques (virtual screening and profiling) to produce vast chemogenomic spaces. He showed several classification schemes for both chemical and biological entities and how this is important to facilitate extraction of knowledge from stored data. For biological entities, he presented unified classification schemes based on unique digit codes, illustrating their use for enzymes and nuclear receptors. For chemical entities, he presented a hierarchical classification scheme for chemical structures, based on the molecular equivalence number (MEQNUM) algorithm. This method uses graph chemical identifiers for different levels of description of molecules (scaffold, sidechains, links, ring systems, and rinks) to derive a unique chemical structure code. This classification scheme is very useful for storing data in databases and can enable filling of annotation gaps in the chemogenomic space.

Grace Patlewicz (ECB) introduced the plenary discussion. Using some open questions, she led the discussion on what might be the different steps in a process map for developing chemical categories. The discussion centred on endpoints of high priority within REACH, including skin sensitisation, mutagencity, carcinogenicity, endocrine disruption and reprotoxicity. The first three endpoints are perhaps better understood in terms of their "mechanisms" or at least there is more toxicity data associated with them that enables associations between chemical structure and effect to be made. For example there is a reasonable amount of public information available for mutagenicity and carcinogenicity from the Carcinogenicity Potency DataBase (CPDB), or US National Toxicology Program (NTP), whereas reprotoxicity data is substantially more limited. The suggestion was that knowledge about these endpoints (from toxicologists) could be formulated into simple structural rules, either by using statistical techniques on the available public datasets and cross checking the output with human experts or by interrogating the experts themselves and encoding their knowledge into a computer program. If data was more limited, surrogate assays could be promising tools in formulating mechanistic hypotheses e.g. the information derived from a peptide binding assay may provide sufficient information to enable some mechanistic information to be derived that can help in the formulation of groupings for skin sensitization. Additionally, metabolism information (using data derived from pharmacologists to determine which chemicals are activated, glucuronidated, sulphonated etc) could be used to understand more about the inherent behaviour of chemicals in order to formulate groupings. Often although mechanistic insights are strong, toxicity cannot be confidently predicted from the structure because the chemistry cannot be predicted. In such cases an "in chemico" approach involving, as appropriate, identification of reaction products, measurement of rate constants and if necessary investigation of oxidation chemistry, can make confident prediction possible, using mechanism based QSAR and Intelligent read-across (i.e. based on comparisons within the same mechanistic applicability domain).

Day 2 – Thresholds of Toxicological Concern

Grace Patlewicz (ECB) summarised the discussions carried out on the first day on chemical category formation.

Andrew McDougal (FDA) was unable to participate in person, but he provided a recorded presentation on how TTCs are applied within the US FDA's Office of Food Additive Safety (OFAS). He started by defining the concept of TTC and how it is used as a prioritisation tool within the FDA. He introduced the TOR (threshold of regulation) concept and explained that the Gold (CPDB) database had been used to define the TOR. He outlined current strategies for refining the TOR, such as using structural classes to identify chemicals of higher concern as well as the use of genetic assays that could lower the risk of carcinogenicity. A combination of the Ames test, mouse lymphoma assay and chromosome aberration assay helped to lower the incidence of carcinogens. An OFAS perspective on chemical similarity was provided – focussing on the (Q)SAR tools used, as well as current efforts to organise historical data into structure searchable databases. Following the recorded presentation, Andrew dialled in from the US to take any questions.

Ian Munro (CANTOX Health Sciences International) presented the concepts and assumptions underpinning TTC. He provided an extensive history of TTC and its evolution from the sixties to the present time. He presented an analysis of the threshold values for the carcinogenic compounds in the Gold database, and for non-carcinogenic endpoints. He also presented the Cramer classification tree as a means of classifying substances into one of 3 structural classes which could be used to define different human exposure thresholds. Finally he illustrated how these thresholds have been applied in the safety evaluation of flavouring ingredients by JECFA, an international expert scientific committee administered jointly by the Food and Agriculture Organization of the United Nations (FAO) and the World Health Organization (WHO).

Maria Wallén (Swedish Chemicals Inspectorate) presented a concise literature review and summary of different TTC approaches that had been carried out by KeMI. In particular she highlighted the advantages, limitations, and uncertainties of this approaches.

Bob Safford (SEAC, Unilever) presented current in house work being undertaken in the area of TTC. He presented the TTC as a useful approach in cases of low consumer exposure; such as contaminant incidents, indirect food additives or flavour components in food. Given that the premise of TTC is that 20% of chemicals are carcinogenic, he discussed whether the use of additional information (in silico, in vitro) could lower the incidence of carcinogens. Using the Gold (CPDB) dataset as a starting point, he used the Cramer classification scheme implemented in ToxTree to classify the chemicals into one of three classes. DEREK was used to identify any structural alerts for mutagenicity and carcinogenicity and Ames or mouse lymphoma data (MLA) was taken from the literature. Each piece of information helped to lower the incidence of carcinogens but the MLA was the most effective. DEREK was comparable to the Ames test in reducing the incidence of carcinogens whereas the Cramer classification scheme was found to be over conservative.

Nina Jeliazkova (IDEA Consult Ltd.) gave an overview of the Cramer scheme and demonstrated how this had been encoded into a new piece of software called toxTree (Version 1). She outlined some of the challenges she had

encountered in building the software and approaches to resolve these. She also gave a demonstration of the software, showing how easy it was to process one or many structures and how to view the results generated. The software development was funded by ECB and the application has now been made available as a free download from the ECB website (see http://ecb.jrc.it/QSAR/qsar_tools/qsar_tools_toxtree.php for further information including how to download).

Chihae Yang (Leadscope® Inc.) presented an overview of grouping, adapted to the outcomes and discussions of the workshop. She started presenting a classification of grouping methods, from knowledge-based methods including the TTC approach, to supervised and unsupervised methods. She exemplified the different grouping methods implemented in Leadscope software, i.e. expert rules that group chemicals into pre-defined hierarchical classes (more than 27000 fragments), Tanimoto, and Jaccard distance similarity coefficients calculated on fingerprints, unsupervised agglomerative nesting methods, supervised recursive partitioning, recursive partitioning with simulated annealing, new measures being currently developed such as bitset, and the modified Tanimoto coefficient, and analogue (surrogate) based grouping techniques. The decision trees generated by machine learning method can augment the knowledge-based TTC approach. Strategies to sequentially grouping compounds were also presented.

Grace Patlewicz (ECB) introduced the second brainstorming session and led the plenary discussion on the basis of a number of issues and questions that arose from the morning's presentations. Discussion points included what modifications if any should be undertaken for the Cramer classification tool, whether TTC could be applied for other endpoints such as skin sensitisation, and what aspects of TTC could be applied in the context of REACH. It was generally agreed that the TTC concept could be difficult to apply in the context of industrial chemicals, since the necessary exposure information is rarely available, and there can be a complex chain of uses down the supply chain. She summarised some of the consensus conclusions and recommendations and outlined the next steps in drafting a report. The participants were thanked for their attendance and contribution and the workshop was closed with a final coffee break.

Pag. 1 ----> Pag. 35

European Commission

EUR 22657 EN – DG Joint Research Centre, Institute IHCP
Title: Chemical Similarity and Threshold of Toxicological Concern (TTC) Approaches: Report of an ECB Workshop held in Ispra, November 2005
Authors: Grace Patlewicz, Ana Gallegos Saliner, Manuela Pavan, Romualdo Benigni,Andrew Worth, Aynur Aptula, Arianna Bassan, Cecilia Bossa, Agneta Falk-Filipsson, Val Gillet, Nina Jeliazkova, Andrew McDougal, Jordi Mestres, Ian Munro, Tatiana Netzeva, Bob Safford, Brigitte Simon-Hettich, Ivanka Tsakovska, Maria Wallén, Chihae Yang
Luxembourg: Office for Official Publications of the European Communities
2007 – 35 pp. – 21 x 29.7 cm
EUR - Scientific and Technical Research series; **ISSN 1018-5593**

Abstract

There are many national, regional and international programmes – either regulatory or voluntary – to assess the hazards or risks of chemical substances to humans and the environment. The first step in making a hazard assessment of a chemical is to ensure that there is adequate information on each of the endpoints. If adequate information is not available then additional data is needed to complete the dataset for this substance. For reasons of resources and animal welfare, it is important to limit the number of tests that have to be conducted, where this is scientifically justifiable. One approach is to consider closely related chemicals as a group, or chemical category, rather than as individual chemicals. In a category approach, data for chemicals and endpoints that have been already tested are used to estimate the hazard for untested chemicals and endpoints. Categories of chemicals are selected on the basis of similarities in biological activity which is associated with a common underlying mechanism of action.

A homologous series of chemicals exhibiting a coherent trend in biological activity can be rationalised on the basis of a constant change in structure. This type of grouping is relatively straightforward. The challenge lies in identifying the relevant chemical structural and physicochemical characteristics that enable more sophisticated groupings to be made on the basis of similarity in biological activity and hence purported mechanism of action. Linking two chemicals together and rationalising their similarity with reference to one or more endpoints has been very much carried out on an ad hoc basis. Even with larger groups, the process and approach is ad hoc and based on expert judgement. There still appears to be very little guidance about the tools and approaches for grouping chemicals systematically.

In November 2005, the ECB Workshop on Chemical Similarity and Thresholds of Toxicological Concern (TTC) Approaches was convened to identify the available approaches that currently exist to encode similarity and how these can be used to facilitate the grouping of chemicals. This report aims to capture the main themes that were discussed.

In particular, it outlines a number of different approaches that can facilitate the formation of chemical groupings in terms of the context under consideration and the likely information that would be required. Grouping methods were divided into one of four classes – knowledge-based, analogue-based, unsupervised, and supervised. A flowchart was constructed to attempt to capture a possible work flow to highlight where and how these approaches might be best applied.

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

The mission of the JRC is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

Publications Office
Publications.eu.int